The Computational Chemistry End Station (ChemES)

Professor C. David Sherrill, co-PI
School of Chemistry and Biochemistry
Georgia Institute of Technology

Papers Acknowledging ChemES Support:

1. "**Density Fitting and Cholesky Decomposition Approximations in Symmetry-Adapted Perturbation Theory: Implementation and Application to Probe the Nature of π-π Interactions in Linear Acenes**," E. G. Hohenstein and C. D. Sherrill, *J. Chem. Phys.* **132**, 184111 (2010).

2. "**Density Fitting of Intramonomer Correlation Effects in Symmetry-Adapted Perturbation Theory**," E. G. Hohenstein and C. D. Sherrill, *J. Chem. Phys.* **133**, 014101 (2010).

3. "**Efficient Evaluation of Triple Excitations in Symmetry-Adapted Perturbation Theory via MP2 Natural Orbitals**," E. G. Hohenstein and C. D. Sherrill, *J. Chem. Phys.* **133**, 104107 (2010).

4. "**Basis Set Consistent Revision of the S22 Test Set of Noncovalent Interaction Energies**," T. Takatani, E. G. Hohenstein, M. Malagoli, M. S. Marshall, and C. D. Sherrill, *J. Chem. Phys.* **132**, 144104 (2010).

5. "**Formal Estimation of Errors in Computed Absolute Interaction Energies for Protein-ligand Complexes**," J. Faver, M. L. Benson, X. He, B. P. Roberts, B. Wang, M. S. Marshall, M. R. Kennedy, C. D. Sherrill, and K. M. Merz, *J. Chem. Theory Comput.* **7**, 790-797 (2011).

Our project goals were (1) to develop next-generation computational chemistry software targeted for large-scale parallel computers in general, and the ORNL Jaguar machine in particular, and (2) to use the emerging computational chemistry software capabilites of the Jaguar Chemistry End Station (ChemES) to perform significant chemical computations.

Supporting aim (1), we performed significant work improving and extending the software infrastructure of our freely-available, open-source electronic structure package, PSI. This work was primarily performed by postdoc Sahan Thanthiriwatte and graduate student Edward Hohenstein, assisted by research scientist Massimo Malagoli and co-PI Professor David Sherrill. This effort was performed in collaboration with other groups, including the Virginia Tech groups of Professors Daniel Crawford and Edward Valeev (also part of the ChemES project) and several members of the Fritz Schaefer group at the University of Georgia.

After initial work on portability improvements and optimizations for the PSI3 release, it was concluded that true deployment of PSI on massively parallel systems like Jaguar would require a complete overhaul of the PSI software infrastructure. In this regard, a parallel programming workshop organized by PI Prof. Robert Harrison in August of 2009 was extremely helpful in providing training in basic parallel programming techniques and best practices. Together with our collaborators at other institutions, we initiated an ambitious project to develop a completely new release of the PSI package (to be named PSI4) geared toward massively parallel systems. In addition to development of improved software infrastructure, we also worked to extend the capabilities of the PSI package. While the Communicator objects in PSI4 were, we also rolled out shared-memory parallelization in significant parts of the code.

Work on PSI4 directly supported by the subcontract to Georgia Tech includes addition of density fitting to MP2 in PSI, development of a parallel density-fitted MP2 code, and development of density-fitted symmetry-adapted perturbation theory (DF-SAPT), and generalized code for multipole moments. We have coded but not yet debugged density-fitting extensions to the parallel local MP2 code implemented by postdoc Ben Mintz. In addition, summer support was provided to graduate students Bill March and Jim Waters, who parallelized their 3-body (Axilrod-Teller) molecular dynamics code for inclusion in PSI.

The DF-SAPT project was particularly successful, giving us the world's fastest SAPT program (and leading to three publications in Journal of Chemical Physics to date). The initial code is adapted for shared-memory parallelization using OpenMP. Full adaptation to distributed-memory parallelization will require concomitant work on parallelizing the underlying PSI Hartree-Fock code, which unfortunately did not fit within the scope of the 2-year project. However, this work continues in conjunction with our collaborators, supported as feasible by other grants. An entirely new, density-fitted Hartree-Fock code has been written by the Sherrill and Schaefer groups, and it is being parallelized by Ben Mintz (now at ORNL).

Supporting aim (2), in consultation with Edo Apra at ORNL, we utilized the massively parallel CCSD(T) capabilities of NWChem as they became available on Jaguar. Most of this work was performed by graduate student Michael Marshall, assisted in some cases by Massimo Malagoli. Dr. Malagoli added some spin-component scaling capabilities to NWChem to enable some of the desired comparisons.

First, we performed several key CCSD(T) computations on benzene trimers from the crystal structure of benzene. This work is aimed at resolving a discrepancy in the literature about the size of three-body interactions in crystalline benzene. Some authors have claimed this term to be essentially negligible (Sherrill group), whereas others (Podeszwa) have claimed that it can be as large as 10% of the lattice energy. The Jaguar computations indicate that the correct result appears to be about half-way in between. One issue needing to be resolved prior to publication is an analysis of the possible effect of different crystal structures used by the two groups.

Second, realizing that some approximate methods such as DFT-D are beginning to approach errors of just a few tenths of one kcal mol^{-1} for noncovalent interactions, further progress in developing such

methods demands more definitively converged benchmark data. Hence, we performed much more accurate CCSD(T)/CBS extrapolations than reported in the popular S22 test set of Hobza and co-workers. The new test set, designated S22A, revises some of the prior interaction energies by as much as 0.6 kcal mol^{-1}. The largest computations (e.g., adenine-thymine) were only possible using CCSD(T) on Jaguar. The new paper (*J. Chem. Phys.* 2010) has already garnered at least 25 citations in its first year of publication.

Third, along similar lines, we collaborated with the group of Kenneth Merz (University of Florida) to obtain accurate CCSD(T)/CBS benchmark energies for several model intermolecular contacts taken from the crystal struture of indinavir with HIV-II protease. This work was meant to assess, on a per-contact basis, the typical errors associated with force-field methods in drug docking studies. The individual errors are rather appreciable, and hence the total error in force-field binding energies becomes disturbingly large. These results have been published in *J. Chem. Theory Comput.* (2011).